Potential of ALOS2 and NDVI to Estimate Forest Above-Ground Biomass, and Comparison with Lidar-Derived Estimates

Gaia Vaglio Laurin 1,*, Francesco Pirotti 2, Mattia Callegari 3, Qi Chen 4, Giovanni Cuozzo 3, Emanuele Lingua 2, Claudia Notarnicola 3 and Dario Papale 1

1 Department for Innovation in Biological, Agro-Food and Forest Systems (DIBAF), University of Tuscia, 01100 Viterbo, Italy; darpap@unitus.it
2 Dipartimento Territorio e Sistemi Agro-Forestali (TESAF)/Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova, 35020 Legnaro, Italy; francesco.pirotti@unipd.it (F.P.); emanuele.lingua@unipd.it (E.L.)
3 European Academy of Bozen EURAC-Institute for Applied Remote Sensing, Viale Druso 1, 39100 Bolzano, Italy; mattia.callegari@eurac.edu (M.C.); giovanni.cuozzo@eurac.edu (G.C.); claudia.notarnicola@eurac.edu (C.N.)
4 Department of Geography, University of Hawai'i at Manoa, 422 Saunders Hall, 2424 Maile Way, Honolulu, HI 96822, USA; qichen@hawaii.edu

* Correspondence: gaia.vl@unitus.it; Tel.: +39-392-3952-067

Academic Editors: Lalit Kumar, Onisimo Mutanga, Xiaofeng Li and Prasad S. Thenkabail

Received: 7 September 2016; Accepted: 21 December 2016; Published: 29 December 2016

Abstract: Remote sensing supports carbon estimation, allowing the upscaling of field measurements to large extents. Lidar is considered the premier instrument to estimate above ground biomass, but data are expensive and collected on-demand, with limited spatial and temporal coverage. The previous JERS and ALOS SAR satellites data were extensively employed to model forest biomass, with literature suggesting signal saturation at low-moderate biomass values, and an influence of plot size on estimates accuracy. The ALOS2 continuity mission since May 2014 produces data with improved features with respect to the former ALOS, such as increased spatial resolution and reduced revisit time. We used ALOS2 backscatter data, testing also the integration with additional features (SAR textures and NDVI from Landsat 8 data) together with ground truth, to model and map above ground biomass in two mixed forest sites: Tahoe (California) and Asiago (Alps). While texture was useful to improve the model performance, the best model was obtained using joined SAR and NDVI (R2 equal to 0.66). In this model, only a slight saturation was observed, at higher levels than what usually reported in literature for SAR; the trend requires further investigation but the model confirmed the complementarity of optical and SAR datatypes. For comparison purposes, we also generated a biomass map for Asiago using lidar data, and considered a previous lidar-based study for Tahoe; in these areas, the observed R2 were 0.92 for Tahoe and 0.75 for Asiago, respectively. The quantitative comparison of the carbon stocks obtained with the two methods allows discussion of sensor suitability. The range of local variation captured by lidar is higher than those by SAR and NDVI, with the latter showing overestimation. However, this overestimation is very limited for one of the study areas, suggesting that when the purpose is the overall quantification of the stored carbon, especially in areas with high carbon density, satellite data with lower cost and broad coverage can be as effective as lidar.

Keywords: ALOS2; mixed forest; biomass; lidar; NDVI
1. Introduction

Forests have a major role in the exchange of carbon between the land surface and the atmosphere, and can represent both carbon sinks and sources through forest growth and by means of deforestation and degradation [1]. Forests are also rich in biodiversity, with higher diversity levels often associated with higher above ground biomass (AGB) [2]. Forest carbon stock estimates are required for reducing uncertainty in the global carbon budget, as well as for biodiversity conservation purposes and forested area management and planning.

Satellite-based remote sensing is fundamental in forest biomass monitoring, as it can support the extrapolation of local ground forest measurements to large extents [3]. The collection of ground data is time and resource demanding, usually covers only limited areas and accessible locations, and is organized in plots of relatively small areas.

Remote sensing of forest carbon stock can be considered a challenging task as the information recorded by a remote sensing instrument is only indirectly related to carbon. Passive optical instruments do not penetrate the dense forest canopy and do not sense the forest compartment where most of the carbon is stored; usually they saturate at low biomass values, with their use being strongly limited by cloud presence; however examples of successful optical-based AGB estimations are reported, especially using high spatial resolution data [4].

Among active sensors, lidar (light detection and ranging) usually generates highly accurate biomass estimates, thanks to its ability to provide detailed vertical forest structure information and to consequently link the strong relationship between forest height and biomass. Lidar is at present considered the premier instrument to quantify carbon stocks [5,6]. A common and successful method to use lidar data for AGB estimation consists in the extraction of forest height measures from the lidar point cloud, use these metrics with field data to build and validate a regression model, and apply the model to the whole area covered by lidar data [7–10]. However, lidar data are presently collected only through on-demand airborne surveys, and thus available only with limited spatial and temporal coverage.

The relationship between radar backscattering and biomass has been illustrated more than two decades ago [11]. Since then, various studies focused on the retrieval of forest structural features from synthetic aperture radar (SAR) due to several advantages, including: the availability of satellites equipped with different SAR sensors, high spatial and temporal resolution of the datasets, extended and often global coverage, and radar insensitivity to cloud cover [12]. The SAR system frequency strongly influences the backscattering to biomass relationship, with the P-band characterized by major sensitivity due to its greater penetration [13]. While waiting for the future launch of the P-band European Space Agency BIOMASS satellite mission, specifically designed to monitor forests [14], AGB could be estimated using SAR at different frequencies, including L-band SAR data.

Data from the Japanese Earth Resources Satellite (JERS) and the Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR), the former operational until 1998 and the latter until 2011, were extensively employed to estimate AGB in different forest ecosystems [15–21]. Previous literature results suggests that there is a saturation level above which there is a loss of sensitivity between AGB and backscattered signal [22]. This saturation point is influenced by forest type and structure, environmental conditions, as well as sensor characteristics, with the commonly observed saturation points usually occurring between 30 and 100 Mg/ha [13,23–25].

There is a recognized influence of plot size on AGB estimates, with larger size often resulting in better accuracy of estimates. Using plots of reduced size and ALOS PALSAR data, AGB estimates with moderate or even limited accuracy are usually obtained: a positive correlation of backscatter at all polarizations with tropical AGB at 0.25, 0.5, and 1 ha scales was found by [17], but they noted that the spatial variability of forest structure and speckle noise in SAR data contributed equally to degrading the sensitivity of radar to AGB at scales less than 1.0 ha. In a Chinese forest, He et al. [26] observed a poor relationship between SAR backscatter and AGB at plot level, while at stand level a logarithm equation could be used to describe the relationship in different biomass ranges. A strong monotonical statistical dependence between ALOS PALSAR and AGB was found by [27] in savanna
pine woodlands using field data collected at 0.25, 0.5, and 1 ha, and a weak dependence using data at 0.1 ha. However, due to the high amount of resources needed to set up and monitor forest plots, it is difficult to establish large sampling areas and obtain field datasets based on large plots [28]; thus, small plots are much more used than large ones in biomass research and mapping activities.

Texture features extracted from SAR data have proved useful in some studies to improve AGB estimates: using RADARSAT-2 C-band dual-polarization data in a complex subtropical forest, Sarker et al. [29] found the Grey Level Co-occurrence Measure (GLCM; [30]) texture features more effective than the original bands; in [31], they found that the addition of GLCM textures improved a joined Landsat–ALOS PALSAR model for AGB estimation in Iranian forests; and Champion et al. [32] found high correlation between GLCM textures extracted from airborne P-band cross-polarization data and AGB in a French Guyana forests characterized by high carbon density.

Another option to improve AGB estimates based on SAR could be the addition of multispectral optical data, to exploit the response from different regions of the electromagnetic spectrum. In this respect, Deng et al. [33] found beneficial the addition of WorldView-2 to ALOS PALSAR in mountain Chinese forests; both Basuki et al. [34] and Fedrigo et al. [35] improved the ALOS PALSAR-based estimation by integrating Landsat 7 ETM+ data in tropical forests.

The ALOS2 continuity mission was launched on May 2014 and currently produces dual and full polarization data. Even if full polarization data are available only in selected regions, the availability of polarizations ratio from dual-pol data, such as in this study, is recognized as an advantage when using SAR for biomass estimation [33]. ALOS2 improvements with respect to the former ALOS satellite include: the exploitation of a dual receiving antenna allowing to broaden the imaged swath, improved spatial resolution with spotlight (from 1 to 3 m) and strip-map mode (from 3 to 10 m); a reduced revisit time, from 46 to 14 days; and the possibility to provide both left and right looking for fast coverage in emergency cases [36,37]. To the best of our knowledge, only one study has evaluated the performance of this new sensor for AGB estimation [38]. Specifically, the improved spatial resolution could represent an advantage to build regression models using small field plots.

In our study, we aimed at using small plots to develop regression models between remote sensing derived data—SAR and NDVI (Normalized Difference Vegetation Index)—and ground measured variables. We also used a lidar-based regression model, and developed a new one, for results comparison purposes. Limited research quantitatively compared AGB estimates obtained using SAR plus NDVI and lidar data types, which have quite different data acquisition costs. Through this effort, the present research aims at providing useful insights for forest monitoring, focusing on two different study sites that were chosen for the availability of accurate ground and lidar data, and for representing two different types of mixed conifer forest.

2. Materials and Methods

2.1. Field Data

The Asiago study site (Figure 1) is part of the Asiago plateau (Province of Vicenza—Italy), and is located in a karstic plain area on the esalpic range of the north-eastern Alps. This site is divided into two areas: the Boscon southern area has an extent of about 32 km$^2$, while the Verena northern area covers approximately 23 km$^2$. Slopes are quite mild and elevation ranges come from about 1100 to 1300 m a.s.l. and from about 1300 to 1750 m a.s.l. in the northern and southern part respectively. Vegetation is mixed conifer forest, composed mostly by spruce stands (*Picea abies*), with presence of silver fir (*Abies alba*), beech (*Fagus sylvatica*), and larch (*Larix decidua*). In 2012, in the framework of a national research project, 33 circular plots of 19.95 m radius (0.1256 ha) were set up, in which height and diameter at breast height (DBH) were measured for each tree with DBH greater than 5 cm; AGB was calculated according to species-specific allometric models [39,40]. Plots were set up according to a stratified random sampling, with stratification based on height classes.
The Tahoe study site is located on the eastern slope of the Sierra Nevada mountain range and is named as the United States Department of Agriculture Forest Service (USDA-FS) Lake Tahoe Basin Management Unit. The Tahoe site (Figure 1) covers about 936 km$^2$, but in this analysis we considered an area of 786.4 km$^2$ after removal of all water bodies and a small southern area affected by cloud presence in optical data. The elevation is between 1900 to 2500 m a.s.l. and slopes are usually mild, although a stronger variability is present, especially in the eastern part of the area. The major vegetation type in Tahoe is mixed conifer forest including: Jeffrey pine (Pinus jeffreyi), white fir (Abies concolor), California red fir (Abies magnifica), lodgepole pine (Pinus contorta), incense cedar (Calocedrus decurrens), quaking aspen (Populus tremuloides), western white pine (Pinus monticola), sugar pine (Pinus lambertiana), western juniper (Juniperus occidentalis), and mountain hemlock (Tsuga mertensiana). At Lake Tahoe, over 1000 trees were mapped in 2012 for 56 circular plots of 17.6 m radius (0.0973 ha) using a Nikon DTM-322 total station. These plots were initially established through the Multi-Species Inventory and Monitoring project and the Lake Tahoe Urban Biodiversity project. Plot locations were selected using a combination of systematic/grid sampling and stratified random sampling. At each plot, all
trees greater than 2 cm in DBH were measured. Tree measurements include species, DBH, tree height, height to live crown, and tree status (live, dead, unhealthy, or sick) [9,41,42]. Tree AGB was calculated according to the Component Ratio Method, adopted by the USDA-FS since 2012 [43,44].

The 89 plots were screened against vegetation changes that occurred between the 2012 field data collection and the 2014 remote sensing datasets acquisition; additionally, plots in areas that were affected by distortion effects (layover, foreshortening, shadowing) in SAR-processed images were excluded. Vegetation changes were detected using very high resolution orthophotos (<1 m) and additional Google Earth imagery. After the screening procedure, we reduced the ground dataset to 75 plots, 52 from Tahoe and 23 from Asiago. The two areas significantly differ in terms of AGB content: Tahoe’s AGB ranged between 21 and 305 Mg/ha, with an average of 146 Mg/ha and standard deviation of 70 Mg/ha, while Asiago had between 210 and 530 Mg/ha, an average of 358 Mg/ha and standard deviation of 83 Mg/ha.

2.2. Remote Sensing Data

We used four ALOS2 dual-pol stripmap SAR scenes, one for Asiago dated 8 October 2014, and three for the Tahoe site, from 4 September and 2 October (two images) 2014. These scenes were selected for being as temporally close as possible to field data collection, during days without precipitations. Conversion from digital number (DN) to the backscattering coefficient \( \sigma_0 \) (sigma-naught, in decibel or dB) was performed according to the method described in [45]:

\[
\sigma_0 = 10 \log_{10} Q^2 + I^2 + CF - A
\]  

where \( CF \) and \( A \) are constants, respectively -83 dB and 32 dB and \( Q \) and \( I \) are the real and imaginary part of the digital number.

The HH and HV scenes, having a resolution about 7 m in range and 3.5 m in azimuth, were multi-looked (one look in range and two in azimuth) and filtered with a Lee Filter (7 × 7 window size) in order to reduce the speckle noise. The images were then geocoded and radiometrically calibrated by using the 30 m SRTM digital elevation model (DEM) for the SAR scenes acquired in Italy and a 3 m lidar-derived DEM for the scenes acquired in California. The radiometric calibration, which is the correction of the \( \sigma_0 \) coefficient obtained considering a flat terrain assumption with the local incidence angle \( \theta_i \), was computed using the DEM and orbital data. In particular, the terrain calibrated normalized radar cross-section \( \sigma_0^c \) is computed as:

\[
\sigma_0^c = \sigma_0 \frac{\sin \theta_i}{\sin \theta_F}
\]  

where \( \theta_F \) is the local incidence angle under the flat terrain assumption. After geocoding, the final spatial resolution was set to 0.000083 degrees (approximately 9 m). The remote sensing processing was conducted using the SARscape module in ENVI5.0 software (Exelis Visual Information Solutions, Boulder, CO, USA).

Two atmospherically corrected and calibrated Landsat 8 images from September 2014 were downloaded from the Google Earth Engine facility to cover the two study areas. The Normalized Difference Vegetation Index (NDVI) was then computed for each image.

2.3. Lidar Data and Derived AGB Maps

For the Tahoe area, lidar data were acquired surrounding Lake Tahoe from 11 August to 24 August 2010 using two Leica ALS50 Phase II laser systems mounted in a Cessna Caravan 208B. The Leica systems were set to a pulse frequency of 83–105.9 kHz, flight height of 900–1300 m, and scan angle of ±14°. The resulting point density is eight pulses per square meter. The airborne lidar data were processed using the Toolbox for Lidar Data Filtering and Forest Studies (Tiffs) [46] to extract the lidar metrics within each plot. A biomass prediction map was then calculated for the area, using the lidar quadratic
mean height (QMH) selected with stepwise procedure as input in a power model \(\text{AGB} = a \times H_{\text{qm}}^b\); where \(a\) and \(b\) are coefficients and \(H_{\text{qm}}\) is the QMH of all returns, trained and validated (five-fold cross validation) with the original 56 field plots, at a 31 m spatial resolution in agreement with the plot size; the datasets and models are also documented by [41].

For the Asiago area, lidar data were collected on 5 June 2012, using a helicopter and an Optech ALTM 3100 sensor with a scan angle \(\pm 29^\circ\) and a scan frequency of 100 kHz. With a relative flight height of 200–725 m, the resulting average point density was 10–12 pulses per square meter [47]. A biomass prediction map at 31 m resolution was calculated, fitting the same model form used in Tahoe with the QMH lidar metric, selected by stepwise procedure, and trained and validated (leave-one-out cross validation) with the original 33 field plots; the full set of available plots were used because they were collected at the time of lidar survey. We used LOO validation for the Asiago site, with respect to the five-fold cross validation used in Tahoe, due to the fewer plots available in Asiago that would have resulted in a limited number of folds and, consequently, less effective validation. However, unreported tests indicated non-significant difference in results when using five-fold for Asiago.

2.4. Data Analysis

The SAR data for Tahoe, acquired in two different dates, were first masked to exclude water pixels and then normalized. For Asiago and Tahoe areas, all the pixels having >70% of their area inside the plot, thus the majority, were extracted from the ALOS2 scenes and weighted according to the percentage of area included into the plot. A set of basic statistics per plot were computed: mean, standard deviation, minimum and maximum backscattering per HH and HV polarizations, and their difference and sum. Two Tahoe plots were covered by overlapping scenes: values from both scenes were extracted and used to generate the statistics.

Similarly to the SAR data, Landsat 8 values from pixels inside the plot for >70% were extracted by weighted average. Preliminary tests suggested that NDVI is not inferior to single bands for forest biomass estimation, and thus the index was retained for the estimation phase. The following eight GLCM texture features were generated for each NDVI, HH, and HV scenes: mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation [30]. A 64-bit quantization level and three different window sizes—3 \(\times\) 3, 5 \(\times\) 5, and 7 \(\times\) 7—were used. The values from different offsets \((0, -1; -1, -1; 1, 0; 0, 1)\) were averaged assuming non-directional effects; then pixels included >70% inside plot area and their weighted average was extracted.

We tested a set of progressively more complex inputs, performing stepwise regression for feature selection and linear regression (or multiple linear regression when multiple inputs were selected) for each set of inputs. The results, validated with leave-one-out (LOO) and 10-fold cross validation, applied at plot level and with plots from both areas, indicate the selected inputs and the accuracy of the model based on those inputs. The inputs were:

1. the SAR statistics (minimum, maximum, mean, standard deviation per HH and HV; HH and HV sum and difference; total 10 inputs);
2. the SAR statistics plus SAR-GLCM textures per HH and HV (total 26 inputs);
3. the SAR statistics plus NDVI and NDVI-GLCM textures (total 19 inputs);
4. SAR HH + HV selected by Test 1 plus the SAR-GLCM texture type selected by Test 2 and the NDVI feature type selected by Test 3 (totaling three inputs).

Applying the best model derived from the aforementioned tests, after resampling the NDVI texture to the same spatial resolution of the SAR data, two AGB prediction maps were generated for the Tahoe and Asiago area. The AGB maps were masked to ensure full overlapping in area extent with the corresponding lidar-derived AGB maps, and resampled to meet their spatial resolution. Finally, the AGB maps derived from multispectral and SAR inputs were compared to those generated using lidar data. Considering that for one of the study areas the AGB model was previously developed using the stepwise approach [40], for comparison purposes we preferred to maintain this method...
over more complex statistical ones. All the data analyses were conducted using MATLAB [48] and R [49] software.

3. Results

With the aim of comparing the different tested datasets, we first introduce the results obtained by lidar: the AGB map for Tahoe (Figure 2, left), based on a power regression model using QMH as input, was characterized by a $R^2$ of 0.92 and RMSE of 23.6 Mg/ha, obtained using field values and five-fold cross validation. The Asiago AGB map, realized with the same model form and input feature, was characterized by a $R^2$ of 0.75 and RMSE of 62.3 Mg/ha (Figure 2, right), obtained using field values and leave-one-out validation.

For the other SAR and NDVI combined datasets, the inputs for the different tests were selected via stepwise procedure. Cross-validated accuracies for each test allowed selection of the best model for AGB estimation; results are presented in Table 1. In the first (a) test based on SAR data, we used as stepwise regression inputs the statistics derived from HH and HV channels averaged at plot level; the stepwise procedure selected the sum of HH and HV backscattering values, and the linear regression $R^2$ resulted equal to 0.59 with RMSE equal to 78.76 Mg/ha with LOO validation, and 0.59 and 78.33 Mg/ha with 10-fold validation.

Table 1. Results of the tests conducted with different SAR and optical inputs, using stepwise selection and multiple linear regression, validated with ground truth. RMSE is expressed both in Mg/ha and as percentage.

<table>
<thead>
<tr>
<th>Inputs for Tests</th>
<th>Selected via Stepwise Criteria</th>
<th>$R^2$ LOO</th>
<th>RMSE LOO (Mg/ha)</th>
<th>$R^2$ 10-Fold</th>
<th>RMSE 10-Fold (Mg/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a). SAR HH and HV various statistics</td>
<td>HH + HV</td>
<td>0.59</td>
<td>78.76 (0.15%)</td>
<td>0.59</td>
<td>78.33 (0.15%)</td>
</tr>
<tr>
<td>(b). SAR HH and HV various statistics + GLCM HH and HV textures</td>
<td>HH + HV</td>
<td>0.65</td>
<td>71.95 (0.14%)</td>
<td>0.65</td>
<td>72.09 (0.14%)</td>
</tr>
<tr>
<td>(c). SAR HH and HV various statistics + NDVI + GLCM NDVI textures</td>
<td>HH + HV</td>
<td>0.66</td>
<td>71.62 (0.14%)</td>
<td>0.66</td>
<td>71.59 (0.14%)</td>
</tr>
<tr>
<td>(d). SAR HH and HV various statistics + 5 × 5 HHmean + 5 × 5 NDVImean</td>
<td>HH + HV</td>
<td>0.66</td>
<td>71.62 (0.14%)</td>
<td>0.66</td>
<td>71.59 (0.14%)</td>
</tr>
</tbody>
</table>
In the second (b) test, we used as input the SAR statistics and the different SAR GLCM features, using textures calculated with a different window size each time. The stepwise procedure always selected the HH + HV sum and the mean texture derived from HH SAR polarization, but the model using the 5 × 5 mean SAR texture performed better than those using textures calculated with the other window sizes; the multiple linear validated regression showed a R^2 of 0.65 (both with LOO and 10-fold) and a RMSE of 71.95 (LOO) and 72.09 (10-fold) Mg/ha.

For the third (c) test, we used as input the SAR statistics, the NDVI and the GLCM NDVI textures, using each time textures calculated with different window size. The NDVI in the two areas showed different averaged values. Namely, 0.52 for Tahoe and 0.69 for Asiago. The stepwise procedure always selected the HH + HV backscattering and the mean NDVI texture, but the model using the 5 × 5 mean NDVI texture performed better than those using NDVI textures calculated with the other window sizes; R^2 for multiple linear regression was equal to 0.66 for both LOO and 10-fold validation, and RMSE to 71.62 (LOO) and 71.59 (10-fold) Mg/ha.

In the fourth (d) test, we used as input the 5 × 5 window mean texture from HH SAR data selected in (b), the 5 × 5 window mean NDVI from Landsat 8 data selected in (c), and the HH + HV backscattering value selected in (a); the stepwise procedure selected only HH + HV and the 5 × 5 window mean NDVI texture, with multiple linear regression having a R^2 (coefficient of determination) and a RMSE (root mean square error) equal to those obtained with test (c).

Among test results, we considered as the best the one produced by the fourth (d) test, having the same results of (c) but being based on much reduced number of input data; selected inputs were HH + HV (dB sum) and the 5 × 5 mean NDVI texture, with the related equation and scatterplot of the predicted vs. observed AGB values presented in Figure 3.

![Figure 3](image-url)  
**Figure 3.** Scatterplot of the observed and predicted AGB values using a multiple linear regression model with HH + HV backscattering and 5 × 5 mean NDVI GLCM texture as inputs.

To visualize the dynamic range of the SAR input, the scatterplot AGB and HH + HV backscattering values are presented in Figure 4. The backscatter range is included between approximately −11 and −30 dB, with saturation occurring approximately over 350 Mg/ha.

The statistical comparison of the AGB maps produced with different datasets, namely the multispectral and SAR dataset and the lidar one, are summarized in Table 2. The maps generated
applying the model that uses SAR and NDVI as inputs are presented in Figure 5; Figure 6 illustrates
the difference in the AGB distributions of the two maps, and Figure 7 shows the difference of
the SAR + NDVI AGB estimate with respect to the lidar map AGB intervals. Based on Table 2,
we calculated the difference of total AGB stored in Asiago according to the SAR plus NDVI and lidar
map, which is equal to $1.00 \times 10^5$ and represents 6.4\% of the AGB amount estimated by lidar. Similarly,
for Tahoe this difference is equal to $2.3 \times 10^6$ and corresponds to 23.7\% of the lidar-based estimated
AGB. The number of 31 $\times$ 31 m pixels in the maps are 54,299 for Asiago and 818,315 for Tahoe.

Table 2. Comparison of the lidar-derived and the SAR + NDVI derived AGB maps for the two study
areas.

<table>
<thead>
<tr>
<th></th>
<th>Asiago SAR + NDVI</th>
<th>Asiago Lidar</th>
<th>Tahoe SAR + NDVI</th>
<th>Tahoe Lidar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean AGB (Mg/ha)</td>
<td>321</td>
<td>301</td>
<td>153</td>
<td>117</td>
</tr>
<tr>
<td>Standard deviation of AGB (Mg/ha)</td>
<td>49</td>
<td>117</td>
<td>64</td>
<td>94</td>
</tr>
<tr>
<td>Total AGB (Mg)</td>
<td>$1.67 \times 10^6$</td>
<td>$1.57 \times 10^6$</td>
<td>$1.20 \times 10^7$</td>
<td>$0.97 \times 10^7$</td>
</tr>
</tbody>
</table>

Figure 4. Scatterplot of AGB and HH + HV backscattering values.

Figure 5. SAR + NDVI based map for Tahoe area (left) and Asiago area (right).
Figure 5. SAR + NDVI based map for Tahoe area (left) and Asiago area (right).

Figure 6. Distribution of AGB values in SAR plus NDVI and lidar-based maps. Boxes represent upper (75%) and lower (25%) quartiles; vertical segments out of boxes are minimum and maximum values; boxes are divided by the median; diamonds represent the mean.

Figure 7. Difference per AGB intervals of the SAR + NDVI map with respect to lidar map. Boxes represent upper (75%) and lower (25%) quartiles; vertical segments out of boxes are minimum and maximum values; boxes are divided by the median.

Table 2. Comparison of the lidar-derived and the SAR + NDVI derived AGB maps for the two study areas.

<table>
<thead>
<tr>
<th></th>
<th>Asiago SAR + NDVI</th>
<th>Asiago Lidar</th>
<th>Tahoe SAR + NDVI</th>
<th>Tahoe Lidar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean AGB (Mg/ha)</td>
<td>321</td>
<td>301</td>
<td>153</td>
<td>117</td>
</tr>
<tr>
<td>Standard deviation of AGB (Mg/ha)</td>
<td>49</td>
<td>117</td>
<td>64</td>
<td>94</td>
</tr>
<tr>
<td>Total AGB (Mg)</td>
<td>$1.67 \times 10^6$</td>
<td>$1.57 \times 10^6$</td>
<td>$1.20 \times 10^7$</td>
<td>$0.97 \times 10^7$</td>
</tr>
</tbody>
</table>
4. Discussion

The lidar-derived AGB maps are characterized by high (Asiago) to very high (Tahoe) accuracies, with differences expected for operating distinct instruments over different sites. For Tahoe, the used linear model was developed and published by Chen [41]. For Asiago, the same model form was adopted as preliminary unreported tests indicated it as the most accurate option; as in Tahoe, the QMH resulted to be the most useful regression input. Both models explain the majority of the AGB variability in the areas, but they are characterized by differences in accuracy. High variability in the accuracy of lidar-based estimates obtained by different studies has been previously observed [50]. It is known that different sources of error, generated while propagating from the ground tree level to the landscape estimation level, may affect lidar-based AGB predictions [51], and it is reasonable to consider that the impact of these errors differs according to the study site. Even if at both sites mixed forests are present, the sites and the collected data are characterized by relevant differences, such as: the lidar systems used, the aerial and field survey characteristics, the species communities, the tree cover density, the biomass density—with significantly larger values of AGB found in Asiago with respect to Tahoe—and the allometric relationships adopted. All these factors are reported as possible sources of error in AGB estimations [51] and may explain the observed difference in accuracy. Overall, the accuracies are in a high range, adding evidence to the value that lidar has for biomass studies.

With respect to results from SAR tests, the sum of HH and HV backscattering values was always selected as input in stepwise procedure, even if this input performed only slightly better in regression than the single channel cross-polarized backscattering value usually used for AGB estimations. Previous research showed that L-band cross-polarized backscatter is more sensitive to biomass variations, whereas the co-polarized signal better captures differences in forest cover fraction [52]. High resolution imagery used in the initial screening phase, as well the difference in NDVI average values found in the two areas, confirms a difference in forest cover, with Asiago having a higher density of trees. The cover information brought by HH polarization, together with the fact that the sum of the two channels reduces the extreme backscattering values and acts as an additional filter effect, may be the reason for the selection in our models of the HH and HV sum.

The use of SAR GLCM features in test (b) consistently improved the accuracy of the model: the stepwise selected texture was the GLCM mean, which is a statistical feature useful to describe homogeneous regions and that further suppress the noise in the data. The NDVI information added to the SAR data in the third (c) test also improved the accuracy of SAR data alone, slightly more than that done by the SAR mean texture. This indicates a complementarity of the SAR and optical datatypes, as already shown by other researches [31,33–35,53], and underlines the importance of forest cover information especially when considering simultaneously different vegetation communities, as in this study. When—in test (d)—both mean NDVI and mean HH SAR textures are added to the SAR backscattering statistics, only the first input is selected by the stepwise procedure, possibly due to the redundant cover information present in both HH and NDVI datasets. However, the result from using SAR data with texture and SAR + NDVI are not so different; considering the broad and free availability of NDVI data from multiple sensors, as well as the extra resources needed to compute texture (with large increase of features in input) we support the use of NDVI as a way of improving accuracy of SAR estimates.

The AGB variation explained by SAR inputs only (test (a)) is in the range of the results obtained by other studies based on ALOS PALSAR data [53–55]. The best model, having HH + HV and mean NDVI texture as input features, shows an accuracy higher than some results obtained by studies using small field plots [17,27,56]. These results support the view that ALOS2 is a valuable tool for AGB estimation, and that its potential can be boosted by the integration of an optical feature such as NDVI.

The scatterplot of the observed vs. predicted values obtained with the most accurate model (Figure 3) reveals that slight saturation effects are present at the high Asiago biomass values. However our results, also according to Figure 4, do not clearly show the usual 100–150 Mg/ha saturation limit reported in literature [1,57,58]; and it is difficult to understand if at higher AGB levels there is a scarce
Saturation effect or increased error in model. These results from a single research might be caused also by environmental characteristics of the study areas (e.g., soil and forest structure). Several sensor improvements suggest that ALOS2 is a valuable tool, able to characterize biomass in dense forests and offering continuity to ALOS. Some of the improved features to consider, with respect to the previous ALOS, are: higher geolocation accuracy, generating a better correspondence between remote sensing and ground data; the higher spatial resolution, allowing the collection of multiple backscatter information, and reducing noise, even from plots of reduced size; and the improved radiometric accuracy, resulting in finer backscattering response at different AGB levels. For full sensor details see the information available online: https://directory.eoportal.org/web/eoportal/satellite-missions/a/alos-2. Even if additional study is needed to evaluate the use of this sensor for AGB research, and its response in different forests, our research indicates the suitability of the new ALOS2 for AGB estimation, even using small field plots as ground truth.

The availability in our study areas of lidar-derived AGB prediction maps allowed the comparison of the estimates obtained using this on-demand data or satellite data. The evaluation of results obtained with different instruments, which acquire data at very different costs, can be of high relevance for forest and natural resources management, especially in areas where repeated carbon density monitoring is requested.

In terms of total AGB stored in each area, SAR and NDVI data overestimated the total AGB of 6.4% for Asiago and 23.7% for Tahoe respectively, compared to lidar data. The analysis of the errors illustrated in Figure 7 shows that SAR and NDVI moderately overestimates at lower biomass ranges, approximately until 200 Mg/ha, while it underestimates at higher AGB density. The larger overall overestimation found in Tahoe compared to Asiago is possibly due to the sparser tree cover characterizing this site, in which the SAR backscattering is therefore more influenced by the ground signal component. Viergever [59] also reported overestimation modeling biomass with SAR data in sparse savanna woodlands, while [60] suggested that pantropical carbon maps may overestimate AGB in savanna areas. However, underestimation of tree heights has also been reported for lidar-derived tree height models, and attributed to the laser beams missing the tree tops, especially at low laser point densities [61–63], even if there are several examples of AGB models developed with low point density lidar [64,65]. It is possible that the lidar-based AGB maps are slightly influenced by this effect, especially for Tahoe where the sparse tree density may have exacerbated the problem. The distribution of the AGB values in maps produced with the two different systems (Table 2) show that the central measures are not so dissimilar; however, the range of local variation captured by lidar-based maps is higher than in SAR plus NDVI maps, which tend to reduce the AGB variation over the sites, especially at the high AGB Asiago range. This suggests that when the purpose is the overall quantification of the stored carbon in a given area, satellite data which has lower cost and is widely available can be effective, while when precise information on AGB spatial distribution is needed, lidar is a better choice. For Asiago, considering that accuracy of the AGB lidar-based estimate is moderate ($R^2 = 0.75$) and that the overestimation of SAR plus NDVI is quite limited (6.3%), the cost-effectiveness of using lidar for biomass monitoring has to be carefully evaluated.

5. Conclusions

Airborne lidar remains, when possible, the first choice for local studies and fine scale information on biomass spatial distribution. The present research shows the continuity from ALOS to ALOS2 in providing reasonably accurate AGB estimates at a coarse scale, in this case also using small forest plots at considerable biomass densities. Given the reduced number of input features used in modeling, and the large availability of NDVI information vs. the extra resources needed to compute and select SAR textures features, the use of NDVI as integration is advisable. The results also underline the importance of considering the degree of forest cover, similarly to what has been observed in other research [66], which in our study influenced both lidar and SAR + NDVI results.
Considering the larger availability of small plots with respect to larger ones from forest research and inventory efforts, and the increasing number of SAR missions with low-cost data expected in forthcoming years, to map AGB with SAR might be a cost-effective choice not only for large global or country-level analysis but also at smaller scales. When both data are available, lidar-based estimation can represent an accurate baseline, as in this study, with satellite data offering repeated monitoring in time.

Acknowledgments: Gaia Vaglio Laurin and Dario Papale thank the EU for supporting the BACI project funded by the EU’s Horizon 2020 Research and Innovation Program under grant agreement 640176. Francesco Pirotti and Emanuele Linguà thank the NEWFOR “NEW technologies for a better mountain FORest timber mobilization”—contract number 2-3-2-FR, Alpine Space Program Interreg IIIB 2007/2013. Gaia Vaglio Laurin thanks Japan Aerospace Exploration Agency (JAXA) for providing ALOS2 scenes (P.I. 1012) and Google Earth Engine for providing Ladsat 8 imagery.

Author Contributions: Gaia Vaglio Laurin and Francesco Pirotti conceived and designed the experiments; Gaia Vaglio Laurin, Emanuele Linguà, Qi Chen, and Dario Papale performed the experiments; Claudia Notarnicola, Mattia Callegari, Francesco Pirotti, and Gaia Vaglio Laurin analyzed the data; Qi Chen contributed reagents/materials/analysis tools; Gaia Vaglio Laurin wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


40. Tabacchi, G.; Di Cosmo, L.; Gasparini, P.; Morelli, S. Stima del volume e della fitomassa delle principali specie forestali italiane. Equazioni di previsione, tavole del volume e tavole della fitomassa arborea epigea. Consiglio per la Ricerca e la sperimentazione in Agricoltura, Unità di Ricerca per il Moni. In Trento: Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Unità di Ricerca per il Monitoraggio e la Pianificazione Forestale; Consiglio per la Ricerca e la sperimentazione in Agricoltura (CRA): Roma, Italy, 2011. (In Italian)

© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).